

# Final Project Proposal



## Topic:

Gender Detection using Parallel Machine Learning

## Group Members:

- Dingling Ge (001493622)
- Zefan Feng (001422191)

## Keywords:

Gender detection, CNN, Keras, OpenMP, Tensorflow, OpenCV

## Introduction:

Gender detection plays an important role in applications such as human computer interface, face recognition, and face image database management. In this project, we are trying to approach a precise binary classification between biological male and female.

This project supports the following objectives:

- Choose an algorithm to implement gender detection (CNN)
- Optimize performance through parallelism (multi-CPU)
- Experimental support
- Analysis
- Make a conclusion

## Methodology:

Traditional machine learning algorithms such as logistic regression and random forest usually perform poorly in classification tasks due to the variety and complexity of the features of images, so we do not consider using them.

We first consider the artificial neural network (ANN) algorithm among the existing algorithms. Then, we would use one Neural Network structure, Convolutional Neural Network (CNN), which is more suitable for image classification tasks.

To better construct the models, we decided to use TensorFlow, which as one of the most popular libraries for deep learning. TensorFlow can provide most best frameworks and API, like keras, and can run on multiple CPU, GPU or both at the same time, which is convenient for parallel implementation. Therefore, TensorFlow is our first choice. In addition, we chose OpenCV to do the basic operation for image processing because the image opened by OpenCV can truly and lossless display the information of the image, which is at the pixel level and it can more accurately process the image with strong scalability and high customization.

For speed up, in this project, we chose to implement parallelism with OpenMP, which is encapsulated in TensorFlow because parallelism can significantly improve the performance.

## Dataset Description:

- Overview: 2280 cropped faces (1140 male, 1140 female).
- Size: 68MB
- Detail: This is Kaggle's annotated face gender recognition dataset, based on color images of European and American celebrities' faces on the Internet, includes a training set (800 female/male), a verification set (170 female/male), and a test set (170 female/male), with different lighting, glasses and head parts.

## Data Sources:

Download: <https://www.kaggle.com/gmlmrinalini/genderdetectionface>

## Reference:

[1] G. Levi and T. Hassner, "Age and gender classification using convolutional neural networks," 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Boston, MA, 2015, pp. 34-42.